LifeNet Health®
LifeSciences

APPLICATION OF ACTIVITY SCORING SYSTEM TO ANALYZE TRENDS IN PRIMARY HUMAN HEPATOCYTES' DONOR GENOTYPES

Mary Jane Cunningham, PhD

LifeSciences Division, LifeNet Health, Virginia Beach, VA

Abstract # 3708/P368

ABSTRACT

One focus of precision medicine is to use a patient's genotype to modulate a therapeutic drug dose to the patient's own specific metabolism rate. Patients classified as ultra-rapid metabolizers of drug metabolizing enzymes (DME) may require a higher than recommended dose to maintain the drug's efficacious dose while poor metabolizers may require a lower dose to maintain the drug's efficacious dose and avoid adverse effects. In drug discovery and safety testing, the incorporation of primary human cells, such has primary human hepatocytes (PHH), as a more relevant model to human *in vivo* has become more common place. In this investigation, a modified method of activity scoring (AS) first proposed by Gaedigk, A., et.al. [Clin. Pharmacol.Ther. 83(2): 234-242 (2008)] was used to analyze trends in donor genotypes of PHH. PHH from145 donors were sequenced to determine the identity of single nucleotide polymorphisms (SNPs)of 138 DME of which 88 are associated with cytochrome P450 SNPs and 24 are associated with Phase II and transporter enzymes SNPs. The sample population ethnicities are: 69.66% European Caucasian, 15.86% African American, 11.72% Hispanic and 2.76% Asian. The AS was applied as recommended with the homozygote reference allele given a score of 1.0, the heterozygote given the score of 0.5 and the homozygote alternate allele given a score of 0.0. However, in this new variation, the scores were applied across all SNPs and summed in order to convert alphabetic designations to numeric for further analysis by various bioinformatic techniques. The sums varied from 15.0 (correlating with impaired or poor metabolism) to 20.5 (correlating with more extensive metabolism) across all ethnicities. For the SNPs sums for Phase II and transporter enzymes, it was observed that PHH from 12/20 African American donors resulted in higher sulfation rates (>73 pmol/min/million cells) with 7-hydroxycoumarin than the PHH from European Caucasian donors. This new variation of AS shows trends across the sample population as well as within the population and may be used more broadly as a tool for optimizing the experimental designs of investigative toxicology assays.

INTRODUCTION

An important aspect of precision medicine is using a patient's genotype to predict their metabolic response to a drug and thereby personalize their therapeutic regimen. In 2008, an activity score (AS) system was designed to assign a numerical value to a patient's alphabetic allele designation for a SNP.[1] The first application was for the polymorphic gene, CYP2D6, which codes for one of the major cytochrome P450s that metabolizes therapeutic drugs. The gene variant SNPs can lead to various adverse effects, such as by decreasing or increasing a drug's metabolism, and therefore changing the drug's efficacy in an individual.

The current phenotypic classifications for metabolic drug response are: a) ultra-rapid metabolizer (UM), b) normal (formerly extensive) metabolizer (NM), c) intermediate metabolizer (IM) and d) poor metabolizer (PM). A patient classified as a PM takes longer to metabolize the drug than a NM resulting in its systematic accumulation and potential adverse effects. A UM breaks down the therapeutic dose so fast that the drug is subsequently non-efficacious.

The AS are numerical scores based on the star (*) nomenclature (SN) for designating haplotypes. The haplotype designations are derived from sequencing the patient's tissue and may contain a single SNP or have a combination of SNPs which act together. For example, while the CYP2D6*4 haplotype contains rs3892097 or g.1846G>A (the transition of G to A at base pair 1846), the CYP2D6*11 haplotype contains rs201377835 (g.882G>C), rs16947 (g.2851C>T) and rs1132840 (g.4181G>C). In the case of the latter, the current consortium rules state that rs201377835 is the determining SNP for that haplotype.[2] In SN, a non-variant haplotype is given the designation *1. Each allele is given its haplotype designation so that the resulting diplotype shows the designation for both the maternal and paternal allele as in the example, *1/*4.

The AS for CYP2D6 is the most complex ranging from 2.0 to 0. For most other genes in which AS has been applied, NM are designated with a score of 1.0, IM with a score of 0.5, and PM with a score of 0.

While AS is one step closer to helping predict the patient's metabolic response to a drug, it does have several drawbacks:

1) AS is dependent on SN, which is prone to change as more genetic information is discovered and added.
2) Established consortia, which determine drug-drug interactions (DDI), update SN and AS and these scientific bodies may not always agree.[3]
3) SN does not always have clearcut guidelines for assigning nomenclature across a broad range of SNP genotypes for a particular gene.
4) AS has only been applied to some genes but not all and the rules may differ depending on the gene and the specific drug under investigation.

Bioinformatics can help determine trends and correlations within scientific data using various mathematical algorithms. However, these algorithms require numerical values. Therefore, it is essential to convert an alphabetic designation, e.g. genotype, to a numerical value, e.g. AS.

So why not apply a numeric value directly to an allele designation without the necessity of overlaying the number on top of a nomenclature system which is changing? A modified method of scoring is presented here for consideration.

MATERIALS & METHODS

Using the Ion AmpliSeq™ Pharmacogenomics Next Generation sequencing panel, specific genomic regions of each PHH were amplified by PCR and then sequenced on the Ion S5™ sequencer (Thermo Fisher, Waltham, MA). The amplicons are then analyzed using a variant caller to determine specific SNP data (NEXT Molecular Analytics, Chester, VA). The panel consisted of 138 DME SNPs: 88 associated with cytochrome P450 and 24 associated with Phase II and transporter enzymes. The remaining SNPs on the panel were more closely associated with DDI and therefore not included in this analysis..

The genotyping ethnicities and allele frequencies were categorized using the National Center for Biotechnology Information's SNP database (Build 154, release date: 4/21/20). The designations of homozygous reference allele, heterozygote, and homozygous alternate allele were used instead of dominant and recessive to simplify the allele call annotations. The scoring was applied as described in Table 1 based on the major allele frequencies. For each PHH lot, the scores were then totaled across all SNPs for all genes and these sum scores were used in the data analysis. The PCA and regression analysis was performed using Minitab® 18.1 Statistical software (Minitab, Inc., State College, PA, release date: 2/28/17).

Table 1.

Allele Call Designation	Score
Homozygous reference allele	1.0
Heterozygote	0.5
Homozygous alternate allele	0.0

RESULTS

Ethnicity	No. of Lots	Percentage
European Caucasian	101	69.66
African American	23	15.86
Hispanic	17	11.72
Asian, East	2	1.38
Asian, Korean	1	0.69
Asian, South	1	0.69
Total	145	100.00

Table 2. PHH lot ethnicities. *Note:* Asian ethnicities are separated into major groups. The three groups in which the PHH donors fall are: a) Korean (extensive databases for the Korean population), b) South (includes Indian), and c) East (includes Chinese, Japanese).

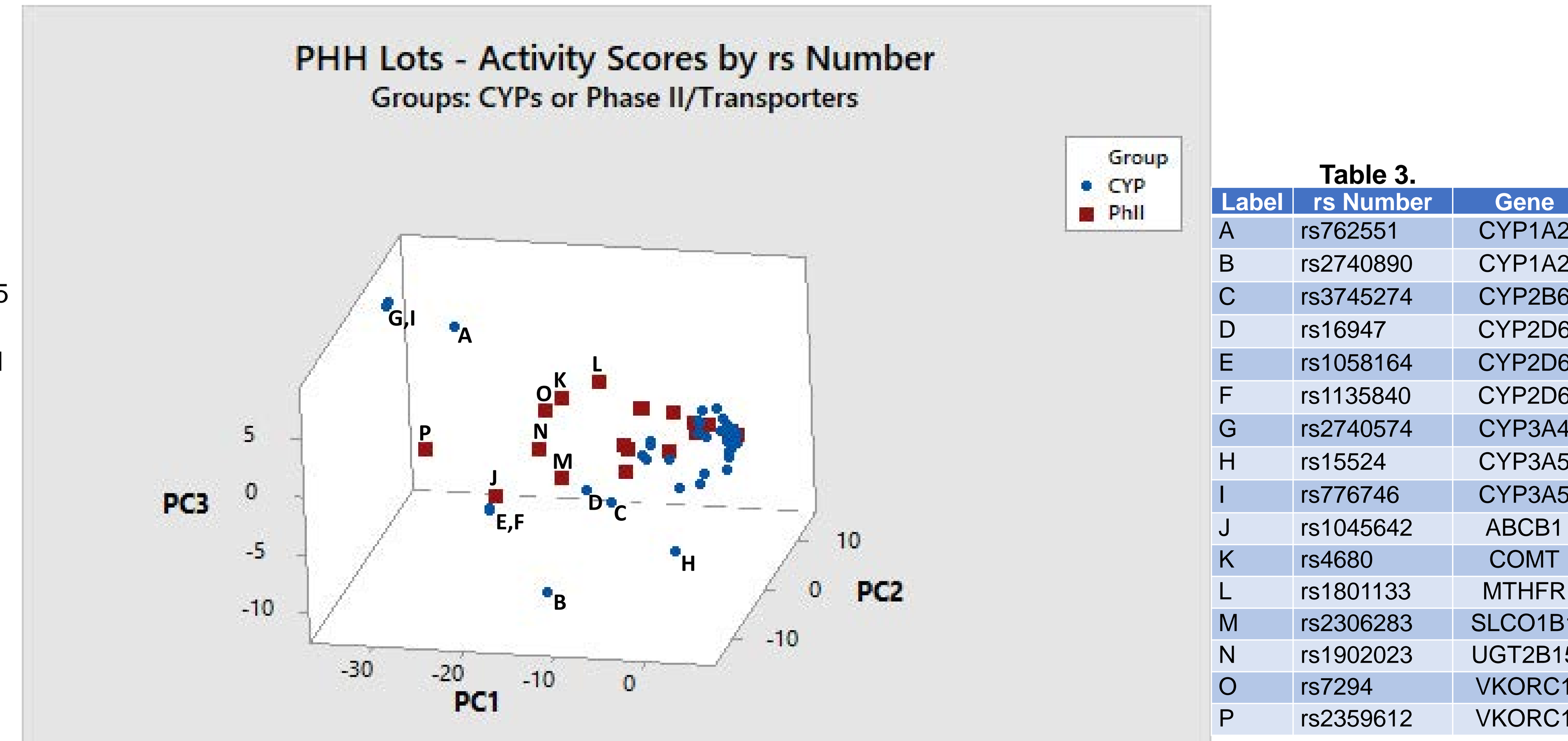


Table 3.

Label	rs Number	Gene
A	rs762551	CYP1A2
B	rs2740890	CYP1A2
C	rs3745274	CYP2B6
D	rs16947	CYP2D6
E	rs1058164	CYP2D6
F	rs1135840	CYP2D6
G	rs2740574	CYP3A4
H	rs15524	CYP3A5
I	rs776746	CYP3A5
J	rs1045642	ABCB1
K	rs4680	COMT
L	rs1801133	MTHFR
M	rs2306283	SLCO1B1
N	rs1902023	UGT2B15
O	rs7294	VKORC1
P	rs2359612	VKORC1

Figure 1. Principal Components Analysis (PCA) showing modified activity scores (sum scores) by rs numbers. The three most statistically significant principal components are shown. The ***most variable*** CYP SNPs and Phase II enzyme/transporter SNPs are listed in Table 3.

RESULTS (CONT.)

To check the diversity of sum scores method, standard alpha diversity indices were calculated.

$$\text{Shannon Index (H)} = -\sum_{i=1}^{S} p_i \ln p_i \text{ ; H = 0.044 (neutral sign)}$$

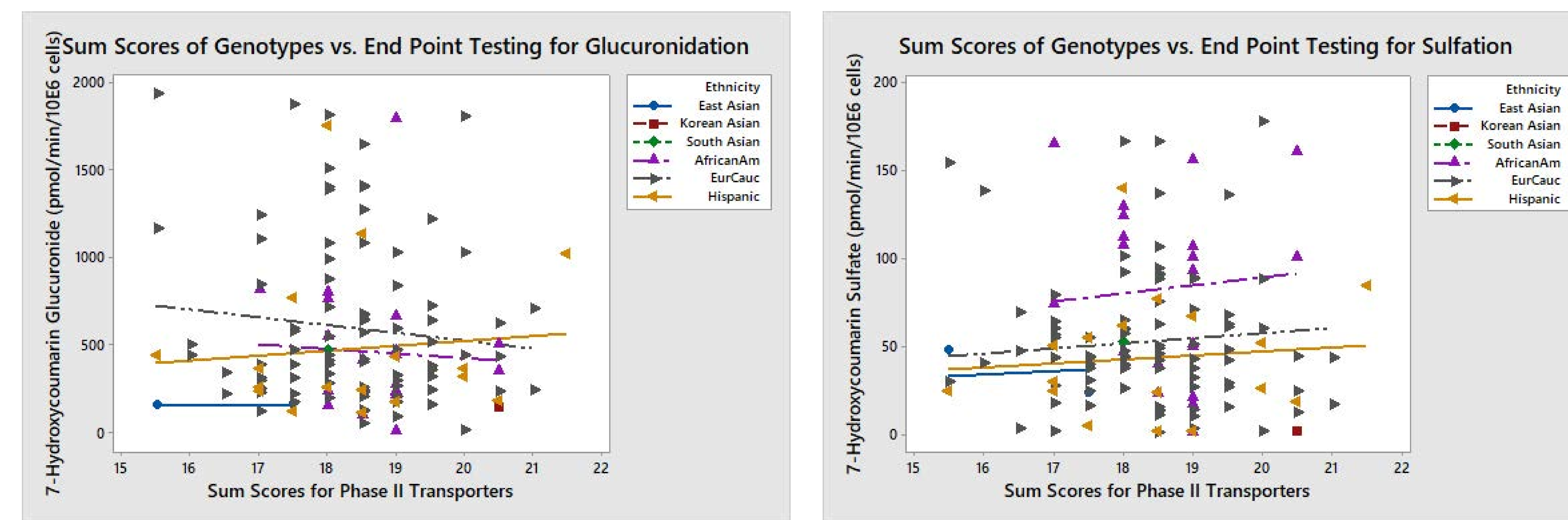
$$\text{Simpson Index (D)} = 1 / \sum_{i=1}^{S} p_i^{2} \text{ ; D = 106.93 with Simpson's Reciprocal Index [1-D] of 0.00935}$$

Traditionally, in both methods, p is the proportion of individuals of one particular species (n) found divided by the total number of individuals found (N). In this case, n = sum scores of one PHH lot and N = the total of all sum scores across all lots. H is an information statistic index and typical values are 1.5 to 3.5. D is a dominance index and would give more weight to dominant sum scores.

The sum scores varied from 15.0 (correlating with impaired or poor metabolism) to 20.5 (correlating with more extensive or normal metabolism). H is 0.044 which falls short of the typical range of a diverse population. D is 106.93 and the Reciprocal Index is 0.00935, both of which show the lack of optimized variability using the sum scores method. The application of these alpha diversity indices could reflect several observations. The donor population for these PHH lots lacks diversity with almost 70% of the donors identifying as European Caucasians. Genotypes indicating variant alleles are naturally present in lesser amounts in the worldwide population. While the sum scores method did increase the diversity of genotype scoring, there are still questions:

a) Does the sum scores method lead to adequate diversity of genotypes?

b) Are the results from these alpha diversity indices just reflecting the typical observations in the worldwide population regarding variant alleles?



Figures 2 and 3. Scatterplots showing comparisons of sum scores for Phase II enzymes and transporters to end point testing values of glucuronide and sulfate forms of 7-hydroxycoumarin, respectively. There is no strong correlation or significance with the formation of glucuronide. However, there is a notable increase in the formation of the sulfate observed in African Americans with 12/20 lots showing >73 pmol/min/million cells of 7-hydroxycoumarin sulfate. The sum scores method has facilitated the observation of a direct comparison of genetic variants data with the formation of a significant DME metabolite.

CONCLUSIONS

- The sum scores method allowed for significant variability to be seen with PCA analysis for genes of the cytochrome P450 family (CYP1A2, CYP2B6, CYP2D6, CYP3A4 and CYP3A5) and well-known transporters (ABCB1, SLCO1B1, UGT2B15) as well as other important genes in drug metabolism (COMT, MTHFR, VKORC1).
- Analysis of the data set by alpha diversity indices showed that this data falls short of a high level of diversity but could be reflecting the actual proportion of genetic variants in the donor PHH population.
- The sum scores method allowed for visualization of genetic variant comparisons with end point testing of DME metabolites.

REFERENCES / ACKNOWLEDGEMENTS

1. Gaedigk A *et al.* (2008) The CYP2D6 activity score: Translating genotype information into a qualitative measure of phenotype. *Clin Pharmacol Ther* **83:** 234-242.
2. Nofziger C *et al.* (2020) PharmVar gene focus: CYP2D6. *Clin Pharmcol Ther* **107**: 154-170.
3. Abdullah-Koolmees H *et al.* (2021) Pharmacogenetics Guidelines: Overview and Comparison of DPWG, CPIC, CPNDS, and RNPGx Guidelines. *Front Pharmacol* **11**: 595219.

I thank Danette Fuentes for assistance with tissue recoveries and Stacey Hanson for assistance with logistics and poster arrangements. Additional thanks gp to the LNH P&L team (Jean Hatfield, Pergentino Balbuena) for isolating the hepatocytes and arranging for their genotyping and NEXT Molecular Analytics for performing the genotyping.

68-90-183.00